Real-time Ensemble Forecasting of Coronal Mass Ejections using the WSA-ENLIL+Cone Model

M. Leila Mays[1,2], A. Taktakishvili[1,2],
A. Pulkkinen[2], P.J. MacNeice[2],
L. Rastaetter[2], D. Odstrcil[2,3], L.K. Jian[2,4],
I.G. Richardson[2,5], J.A. LaSota[6], Y. Zheng[2],
M.M. Kuznetsova[2]
and the CCMC/SWRC team

Email: m.leila.mays@nasa.gov,
aleksandre.taktakishvili-1@nasa.gov,
antti.a.pulkkinen@nasa.gov

[1]CUA, [2]NASA GSFC, [3]GMU, [4]UMD, [5]CRESST, [6]UI

European Space Weather Week 2014

manuscript submitted to Solar Physics

Abstract

Ensemble modeling of coronal mass ejections (CMEs) provides a probabilistic forecast of CME arrival time which includes an estimation of arrival time uncertainty from the spread and distribution of predictions and forecast confidence in the likelihood of CME arrival. The real-time ensemble modeling of CME propagation uses the Wang-Sheeley-Arge (WSA)-ENLIL+Cone model installed at the *Community Coordinated Modeling Center* (CCMC) and executed in real-time at the CCMC/*Space Weather Research Center*. The current implementation of this ensemble modeling method evaluates the sensitivity of WSA-ENLIL+Cone model simulations of CME propagation to initial CME parameters. We discuss the results of real-time ensemble simulations for a total of 35 CME events which occurred between January 2013 - July 2014. For the 17 events where the CME was predicted to arrive at Earth, the mean absolute arrival time prediction error was 12.3 hours, which is comparable to the errors reported in other studies. For predictions of CME arrival at Earth the correct rejection rate is 62%, the false-alarm rate is 38%, the correct alarm ratio is 77%, and false alarm ratio is 23%. The arrival time was within the range of the ensemble arrival predictions for 8 out of 17 events. The Brier Score for CME arrival predictions is 0.15 (where a score of 0 on a range of 0 to 1 is a perfect forecast), which indicates that on average, the predicted probability, or likelihood, of CME arrival is fairly accurate. The reliability of ensemble CME arrival predictions is heavily dependent on the initial distribution of CME input parameters (e.g. speed, direction, and width), particularly the median and spread. Preliminary analysis of the probabilistic forecasts suggests undervariability, indicating that these ensembles do not sample a wide enough spread in CME input parameters. Prediction errors can also arise from ambient model parameters, the accuracy of the solar wind background derived from coronal maps, or other model limitations. Finally, predictions of the Kp geomagnetic index differ from observed values by less than one for 11 out of 17 of the ensembles and Kp prediction errors computed from the mean predicted Kp show a mean absolute error of 1.3.

Ensemble Modeling

Ensemble modeling is used in weather forecasting to quantify prediction uncertainties and determine forecast confidence

- Individual forecasts which comprise an ensemble forecast represent possible scenarios which approximate a probability distribution that reflects forecasting uncertainties.
- Uncertainties can be from initial conditions, observation error, and techniques and models.
- Different forecasts in the ensemble can start from different initial conditions and/or be based on different forecasting models/ procedures.
- Provides a quantitative description of the forecast probability that an event will occur by giving event occurrence predictions as a percentage of ensemble size (probabilistic forecast).
- Conveys the level of uncertainty in a given forecast in contrast to a categorical yes/no forecast. (Categorical forecasts only have two probabilities, zero and one).

Ensemble Modeling with WSA-ENLIL+Cone

The current version of real-time ensemble modeling at the CCMC/SWRC evaluates the sensitivity of CME arrival time predictions from the WSA-ENLIL+Cone model to initial CME parameters.

- Create a set of N CME input parameters with a CME analysis tool (such as StereoCAT). Typically N=36 to 48 provides an adequate spread of input parameters, and this number can be increased if necessary.

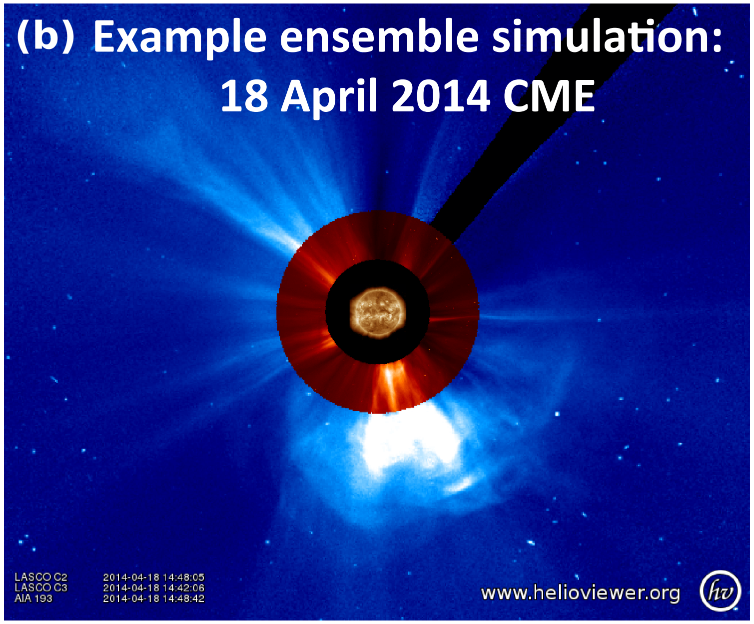
- These are used as input to an ensemble of **N** WSA-ENLIL+Cone model runs.

- This gives an ensemble of **N** profiles of MHD quantities and **N** CME arrival time predictions at locations of interest.

- At Earth, **N** Kp estimates are made using WSA-ENLIL+Cone model plasma parameters as input to the Newell et al. (2007) coupling function for three clock angle scenarios (Θ_C=90°, 135°, and 180°).

- For N=48, a average run takes ~130 min on 24 nodes (4 processors/node) on our development system. We estimate that the same run will take ~80 minutes on the production system (16 processors/node).

(b) Example ensemble simulation: 18 April 2014 CME

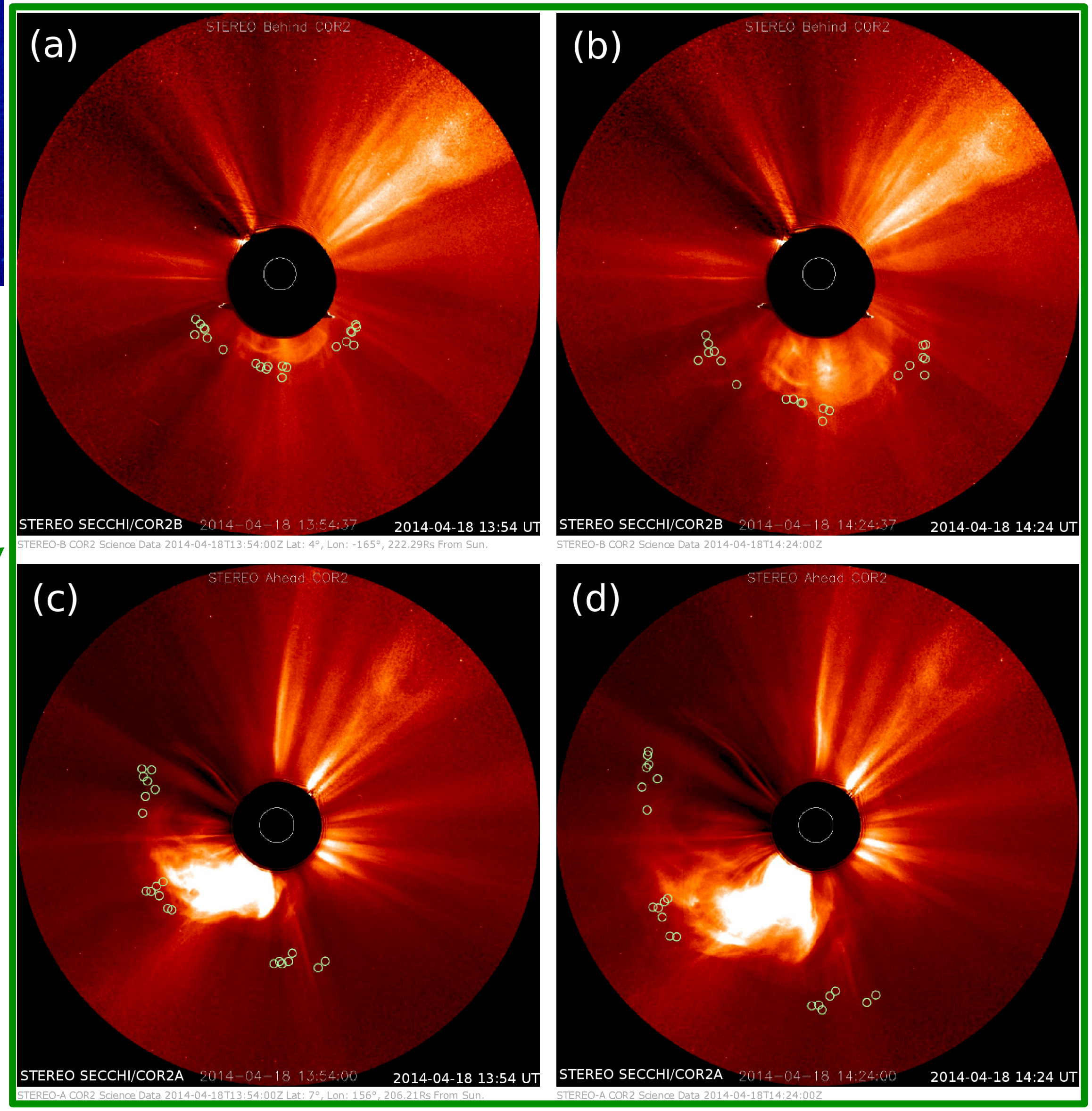


• Earth-directed partial halo CME that was first observed at 13:09 UT on 18 April 2014 by by SECCHI/COR2A. This CME was associated with an M7.3 class solar flare from Active Region 12036 located at S18°W29° with peak at 13:03 UT.

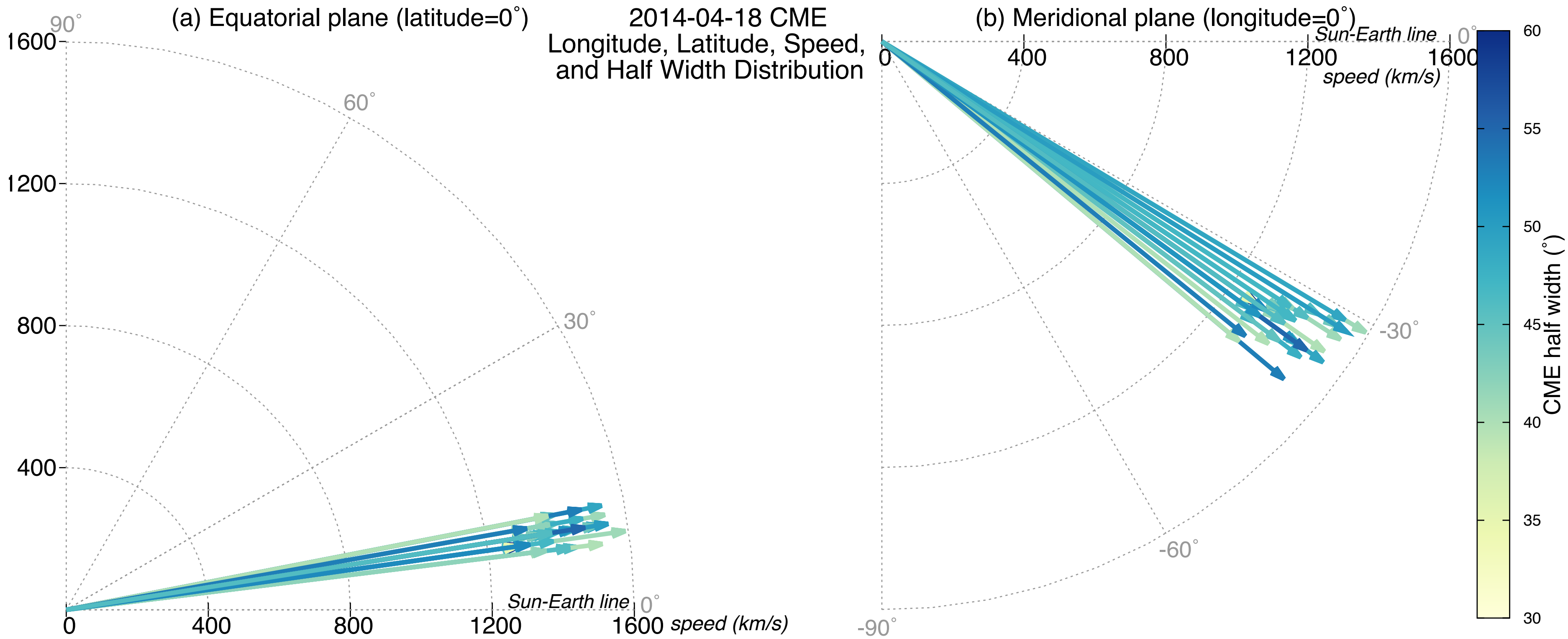
• Eruption and a coronal wave were visible south of the AR in SDO/AIA 193 Å and a nearby filament eruption was visible in AIA 304Å.

•Subsequently starting at 13:35 UT, an increase in solar energetic particle proton flux above 0.1 pfu/MeV observed by GOES Earth orbit.

Ensemble of input CME parameters obtained by measuring the same feature using StereoCAT which employs geometric triangulation techniques. The circles indicate the 6 individual leading edge (circles near the center of the CME front) and width measurements (circles marking the CME edges). The leading edge measurements (central circles) are later combined together to generate 6^2=36 ensemble members.

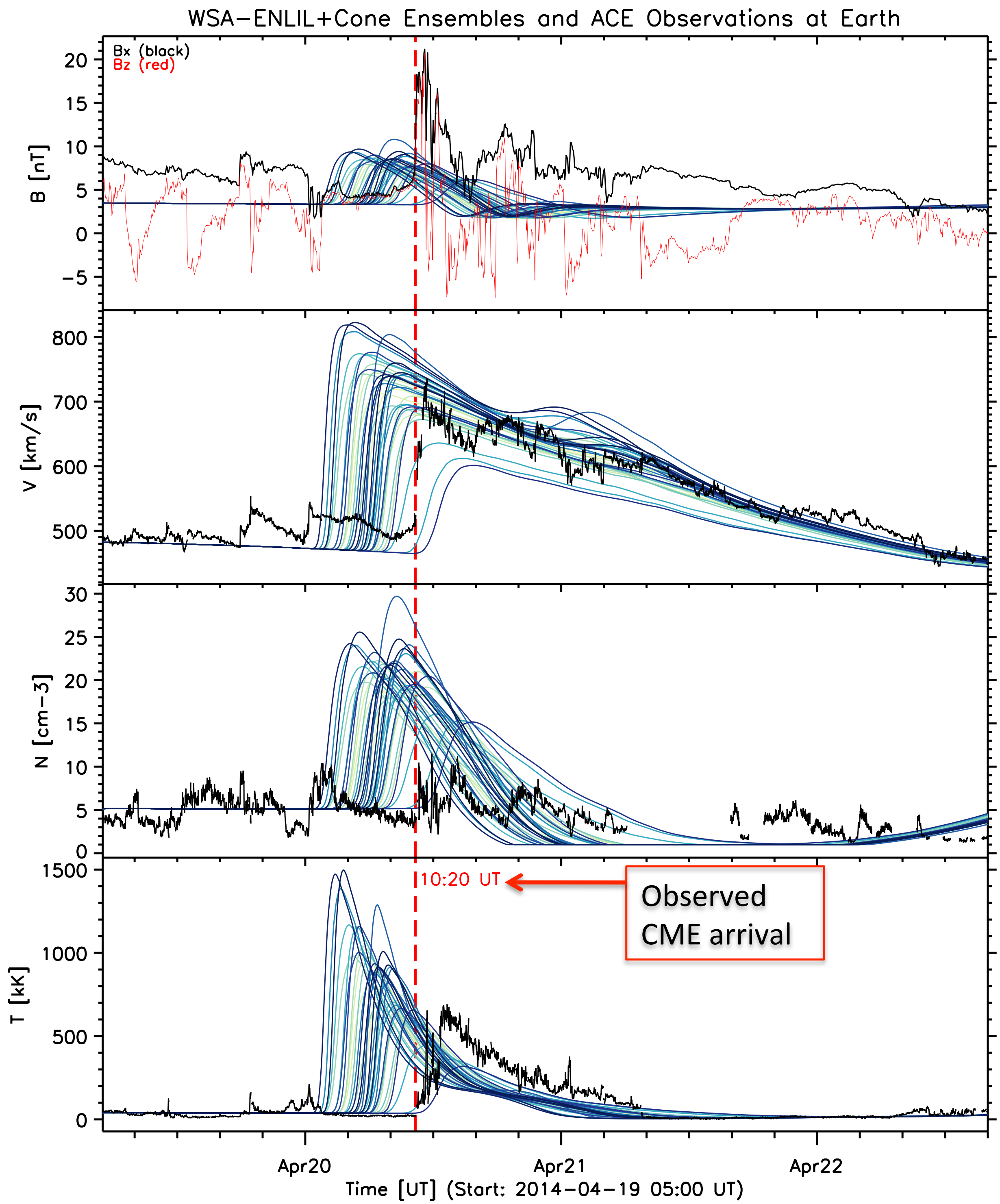


Distribution of the 18 April 2014 CME input parameters



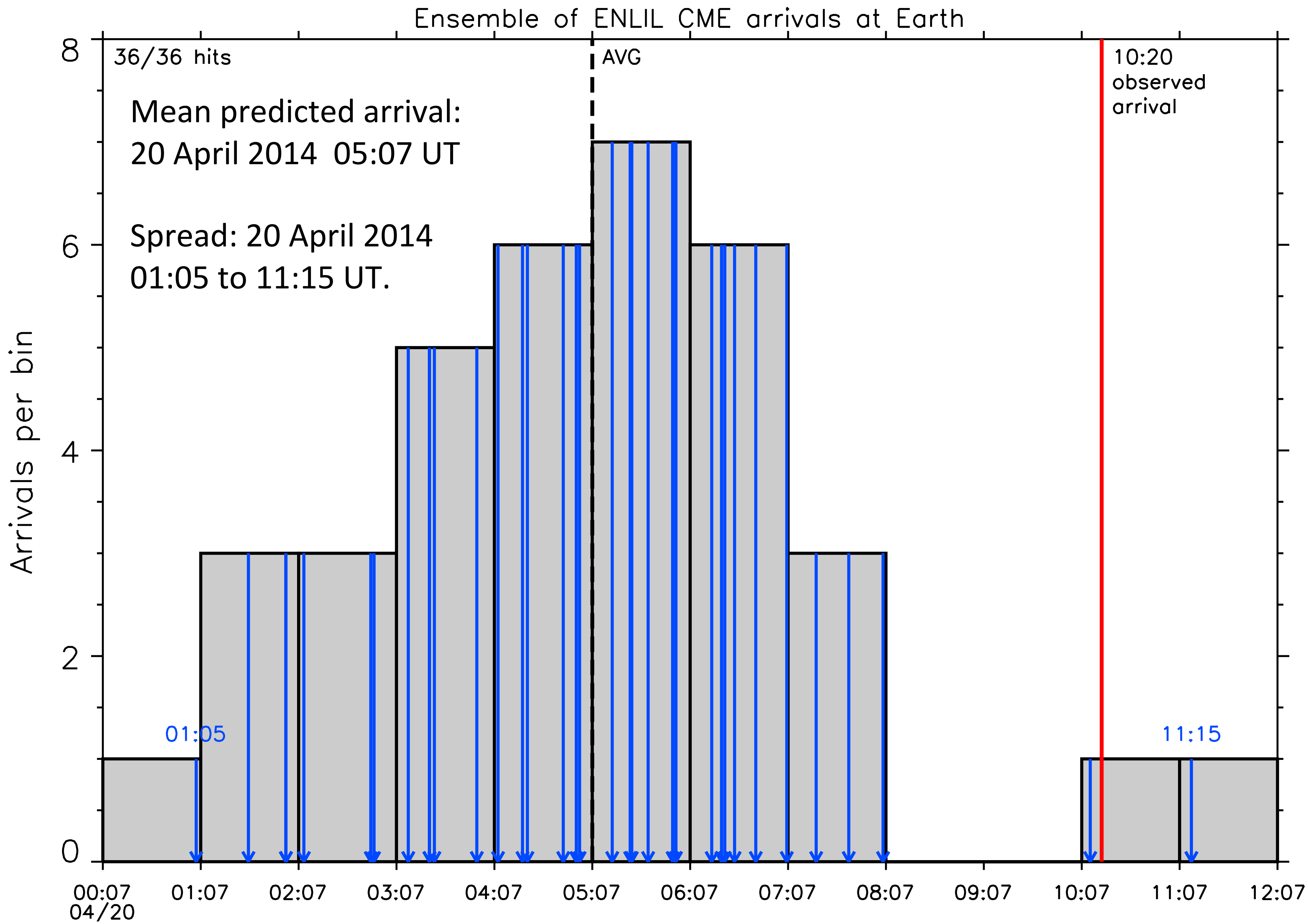
• CME speed vectors in spherical HEEQ coordinates with the grids showing the degrees longitude (a) and latitude (b), and the radial coordinate showing the speed in km/s. The Sun-Earth line is along 0° longitude and latitude. The arrow directions on the grid indicate the CME central longitude and latitude respectively, with CME half width indicated by the color of the vector. The arrow lengths correspond to the CME speed.

• CME propagation directions are clustered between -30° to -40° latitude, and around 10° west of the Sun-Earth line in longitude, while CME speeds range from ~1300 to 1600 km/s. Median CME parameters are: speed of 1394 km/s, direction of 9° longitude, -35° latitude, and a half-width of 46°.



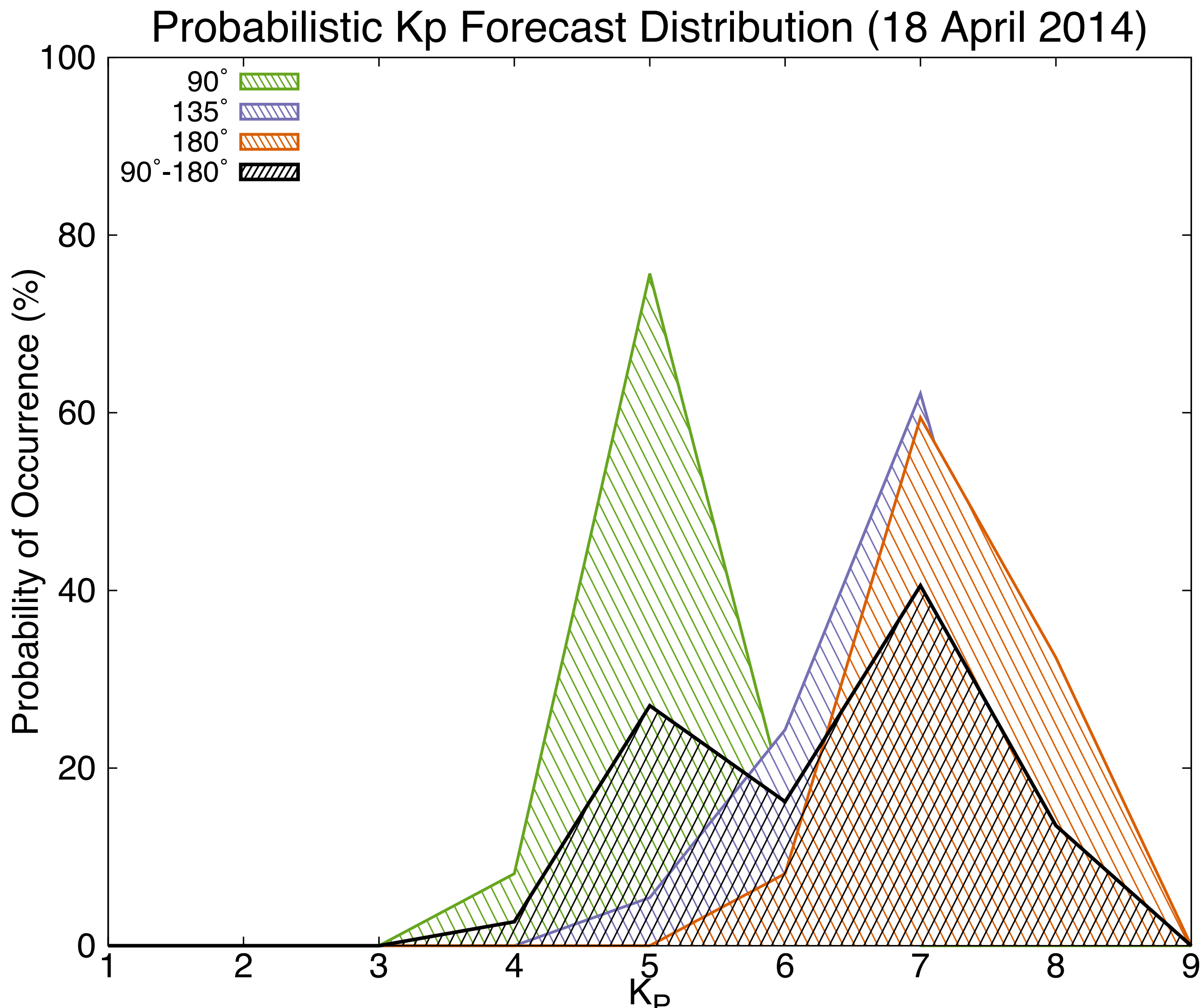
18 April 2014 CME: WSA-ENLIL+Cone modeled magnetic field, velocity, density, and temperature profiles at Earth for 36 ensemble members (color traces), with the observed in-situ L1 observations from ACE (black, red for Bz). Observations show clear signatures of the arrival of an ICME, including a leading shock (abrupt increase in all the solar wind parameters at around 10:20 UT) with enhanced post-shock temperatures, enhanced magnetic field with rotations in direction, and declining solar wind speed.

18 April 2014 CME: Histogram distribution of arrival time predictions at Earth



Shows a normal distribution with 50% of the predicted arrivals within one hour of the mean. The prediction error $\Delta t_{error}=t_{predicted}-t_{observed}$ for the mean predicted CME arrival time was -5.2 hours and the observed arrival time was just within the ensemble predicted spread.

18 April 2014 CME: Distribution of Kp probability forecast



Kp is forecast using ENLIL predicted solar wind quantities at Earth as input to the Newell et al. (2007) coupling function for three clock angle scenarios (Θ_C=90°, 135°, and 180°) and all three angles combined, assuming equal likelihood.

- The observed Kp index reached 5 during the synoptic period 12:00-15:00 UT on 20 April associated with the CME shock arrival.
- 84% of the forecasts fall between Kp = 5 to 7. The most likely forecast is for Kp=7 at 41%, followed by Kp=5 at 27% and Kp=6 at 16% likelihood of occurrence.
- Using the most likely forecast of Kp=7, the prediction error is $Kp_{error} = Kp_{predicted} - Kp_{observed} = 2$ (overprediction)

Summary of Ensemble Simulations for 35 CME Events

Table 1.: Summary of the ensemble simulation results for 35 CME events (January 2013 - June 2014). Columns 1-2: CME onset date and time. Column 3: SOHO LASCO CME Catalog plane-of-sky (POS) speed for 2013. Columns 4-7: median ensemble CME input parameters of v, latitude, longitude (HEEQ), and half-width ($w/2$). Columns 8-10: mean predicted arrival time of all n_{tot} ensemble members, and the spread in arrival times in hours relative to the mean. Column 11: $n_{\mathrm{predicted\ hits}}$, the number of ensemble members predicting that the CME will arrive at Earth out of n_{tot}, the total number of ensemble members. Columns 12-14: actual arrival time observed in-situ, and the observed CME transit time relative to the CME start time. Column 15: prediction error $\Delta t_{\mathrm{err}} = t_{\mathrm{predicted}} - t_{\mathrm{observed}}$ for hits, or CR and FA for correct rejections and false alarms.

CME Onset Date (yyyy-mm-dd)	CME Onset Time (UT)	SOHO v_{POS} (km/s)	Median CME parameters v (km/s)	Lat (°)	Lon (°)	$w/2$ (°)	Mean Predicted Arrival Date (UT)	Time (UT)	Spread (h)	$p_i = n_{\mathrm{hits}}/n_{\mathrm{tot}}$	In-situ Arrival Date (UT)	Time (UT)	Transit Time (h)	Δt_{err} (h)
2013-01-13	07:24	229†	342	1	10	28	2013-01-17	06:30	$^{+5.6}_{-6.2}$	20/48	2013-01-16	23:25‡	....	.....
2013-01-16	19:00	616	750	-26	52	42	2013-01-19	21:33	$^{+11.1}_{-16.3}$	18/48	2013-01-19	16:48	69.8	4.8
2013-02-06	00:24	1851	1460	30	-26	30	2013-02-08	05:37	$^{+8.7}_{-6.8}$	19/48	2013-02-08	03:15‡	....	.....
2013-04-11	07:24	819	1000	0	-15	55	2013-04-13	06:14	$^{+6.1}_{-5.4}$	36/36	2013-04-13	22:13	62.8	-16.0
2013-06-21	03:12	1903	1997	-15	-48	60	2013-06-22	13:02	$^{+5.7}_{-3.6}$	47/48	2013-06-23	03:51	48.7	-14.8
2013-06-30	02:24	410†	386	9	4	34	2013-07-02	20:56	$^{+1.1}_{-0.6}$	4/36	....	......	....	CR
2013-07-16	04:00	639	795	-19	9	19	2013-07-18	20:29	$^{+6.9}_{-6.2}$	28/48	2013-07-18	12:55‡	....	.....
2013-08-02	13:24	443	596	-16	28	13	....	.....	.....	0/24	....	......	....	CR
2013-08-07	18:24	473	570	-25	11	44	2013-08-11	05:03	$^{+6.8}_{-5.9}$	48/48	....	......	....	FA
2013-08-08	23:54	411†	454	-17	14	18	....	.....	.....	0/48	....	......	....	CR
2013-08-30	02:48	884	861	21	-48	59	2013-09-01	08:34	$^{+4.6}_{-5.4}$	46/48	2013-09-02	01:56	71.1	-17.4
2013-09-19	03:36	449†	362	-15	-43	27	....	.....	.....	0/24	....	......	....	CR
2013-09-29	20:40	1164	1000	26	30	66	2013-10-02	04:11	$^{+9.1}_{-10.8}$	36/36	2013-10-02	01:15	52.6	2.9
2013-10-06	14:39	710†	747	1	2	16	2013-10-09	22:10	$^{+10.1}_{-9.8}$	22/24	2013-10-08	19:40	53.0	26.5
2013-10-22	04:24	697	764	51	-10	49	2013-10-25	08:19	$^{+10.2}_{-10.9}$	45/47	....	......	....	FA
2013-12-04	23:12	585	697	41	-9	46	2013-12-07	13:45	$^{+4.1}_{-4.1}$	37/48	....	......	....	FA
2013-12-05*	00:00	623†	651	25	63	31								
2013-12-12	03:36	943	1067	-32	51	50	2013-12-14	18:11	$^{+16.1}_{-13.9}$	36/48	2013-12-15	16:30‡	....	.....
2013-12-12*	06:24	695	694	-52	8	50								
2013-12-29	00:12	296†	682	32	8	47	2014-01-01	02:39	$^{+13.2}_{-9.3}$	48/48	....	......	....	FA
2013-12-29*	05:48	477	495	-33	-58	43				48/48				

* CMEs was simulated together with the CME listed on the previous row as part of a single ensemble.

† 2nd-order plane-of-sky speed at last possible measured height.

‡ In-situ signature could not be unambiguously identified as arrival of CME-related disturbance.

Continued on next page

Table 1.: Continued from previous page

CME Onset		SOHO	Median CME parameters				Mean Predicted Arrival			$p_i =$	In-situ Arrival		Transit	
Date	Time	v_{POS}	v	Lat	Lon	$w/2$	Date	Time	Spread	$n_{\mathrm{hits}}/$	Date	Time	Time	Δt_{err}
(yyyy-mm-dd)	(UT)	(km/s)	(km/s)	(°)	(°)	(°)	(UT)		(h)	n_{tot}	(UT)		(h)	(h)
2014-01-07	18:24		2399	-28	38	64	2014-01-09	00:17	$^{+9.2}_{-6.9}$	48/48	2014-01-09	19:39	49.3	-19.4
2014-01-30	16:24		843	-50	-28	45	2014-02-02	10:10	$^{+11.9}_{-12.3}$	13/24	2014-02-02	23:20	78.9	-13.2
2014-01-31	15:39		718	12	-29	40	2014-02-03	17:20	$^{+9.1}_{-6.0}$	12/12	....	......	....	FA
2014-02-04	01:09		778	-35	21	49	2014-02-06	20:39	$^{+22.8}_{-18.03}$	23/24	2014-02-07	16:28‡	....	.....
2014-02-04*	16:24		560	-34	23	42								
2014-02-12	05:39		740	8	5	59	2014-02-14	23:47	$^{+13.2}_{-8.7}$	48/48	2014-02-15	12:46	79.1	-13.0
2014-02-18	01:25		882	-24	-43	52	2014-02-20	16:29	$^{+17.7}_{-28.7}$	29/36	2014-02-20	02:42	49.3	13.8
2014-02-19	16:00		883	-32	-10	29	2014-02-22	12:20	$^{+13.2}_{-12.6}$	32/36	2014-02-23	06:09	86.2	-17.8
2014-02-25	01:09		1394	-18	-80	78	2014-02-26	22:15	$^{+20.7}_{-11.3}$	40/48	2014-02-27	15:50	62.7	-17.6
2014-03-23	03:48		715	-5	-60	55	2014-03-26	00:58	$^{+11.8}_{-12.6}$	38/48	2014-03-25	19:10	63.4	5.80
2014-03-23	06:36		503	37	-45	34	....	.....	.....	0/12	....	......	....	CR
2014-03-29	18:12		707	36	41	43	2014-04-01	21:30	$^{+1.0}_{-1.0}$	2/36	....	......	....	CR
2014-04-02	13:36		1527	19	-55	51	2014-04-04	19:01	$^{+6.3}_{-10.1}$	14/16	2014-04-05	10:00	68.4	-15.0
2014-04-18	13:09		1394	-35	9	46	2014-04-20	05:07	$^{+6.1}_{-4.0}$	36/36	2014-04-20	10:20	45.2	-5.2
2014-06-04	15:48		580	-40	-28	50	2014-06-07	20:56	$^{+6.8}_{-8.1}$	22/36	2014-06-07	16:12	72.4	4.7
2014-06-10	13:09		980	-9	-89	64	2014-06-12	20:28	$^{+3.1}_{-3.1}$	2/36	....	......	....	CR
2014-06-19	17:12		569	3	-20	44	2014-06-22	16:12	$^{+7.6}_{-5.4}$	12/12	2014-06-22	18:28	73.3	-2.3
2014-06-30	07:24		751	-12	-63	29	2014-07-02	20:32	$^{+7.4}_{-12.3}$	0/36	....	......	....	CR

* CMEs was simulated together with the CME listed on the previous row as part of a single ensemble.

‡ In-situ signature could not be unambiguously identified as arrival of CME-related disturbance.

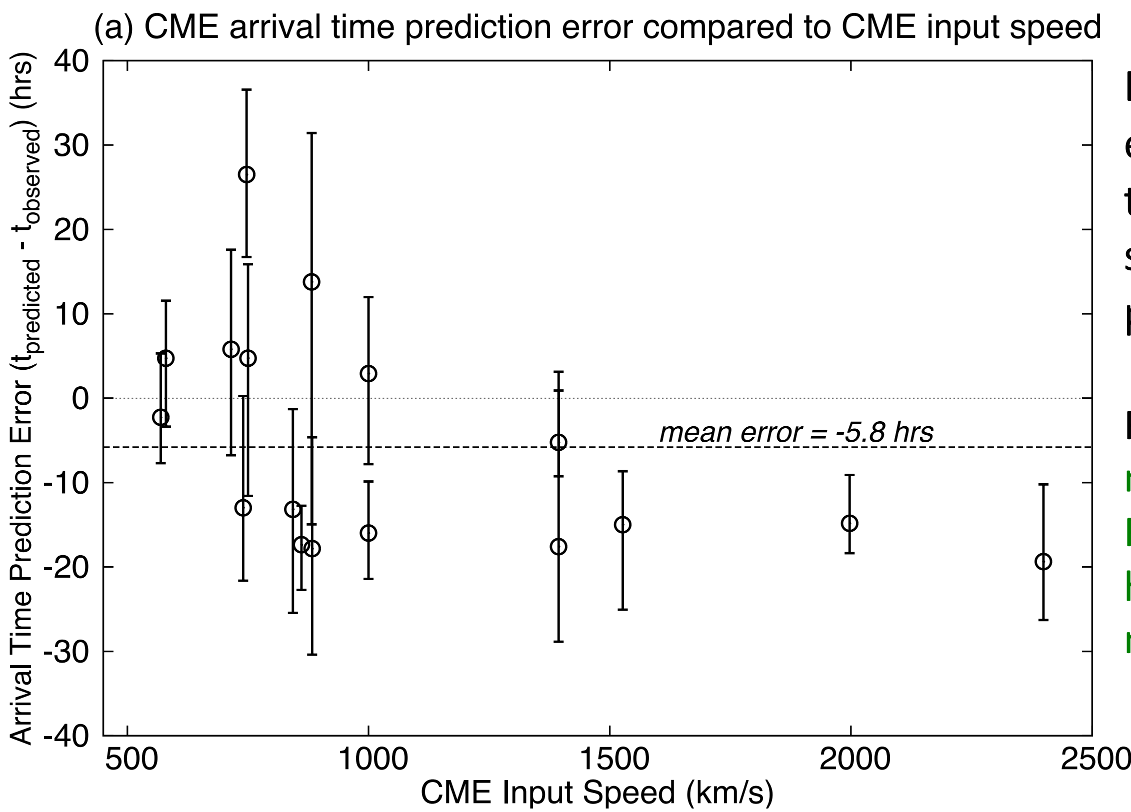
Ensemble forecast verification

Forecast verification was performed for 35 Earth-directed CME events from January 2013 - June 2014 (see Table of events) for which real-time ensemble modeling was carried out by CCMC/SWRC.

I) CME forecast verification

For each ensemble *i*, the forecast consists of a spread of predicted CME arrival times and a forecast probability about the likelihood that the CME will arrive (computed as from the number of predicted hits as a percentage of ensemble size $p_i = n_{\text{predicted hits}}/n_{\text{total}}$).

Ia) CME arrival time prediction forecast verification



Begin with a simple forecast evaluation of CME arrival time, by taking the ensemble mean as a single forecast. Compute the prediction error:

$$\Delta t_{\text{error}} = t_{\text{predicted}} - t_{\text{observed}}$$

For all 17 ensembles containing hits:

mean absolute error (MAE)= 12.3 hrs

Root Mean Square Error (RMSE) = 13.9 hours

mean error (ME) = -5.8 hours (early)

Ib) Likelihood of CME arrival forecast verification

Ensemble modeling provides a probabilistic forecast for the likelihood that the CME will arrive, but begin with simple forecast evaluation by binning the probability p into a categorical yes/no forecast. Categorical forecasts only have two probabilities, zero and one.

• The probability forecast p was binned into two categories: ``yes'' the CME will arrive, and ``no'' the CME will not arrive.

• Correct rejections were identified when the criteria of the forecast probability $p < 15\%$ was met; i.e., that less than 15% of the total predictions in the ensemble indicated CME arrival. Similarly, the inverse criteria is used to identify hits:

Table 2. Forecast performance contingency table for 30 ensembles.

	CME arrival forecast	
Observation	Will occur	Will not occur
Occurs	Hit (17)	Miss (0)
Does not occur	False alarm (5)	Correct rejection (8)

• There were 8 out of 30 correct rejections and 5 false alarms for events that were not observed in-situ, giving a correct rejection and false-alarm rate of 62% (8/13) and 38% (5/13) respectively.

• The correct alarm ratio, defined as the number of hits over the number of hits and false alarms, is 77% (17/22) and the false alarm ratio is 23% (5/22).

Ib) Likelihood of CME arrival forecast verification: Brier Score

Next consider a more nuanced technique to evaluate the probabilistic forecast:

- A method defining the magnitude of probability forecast errors is the Brier Score:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

N = number of events,
p_i = forecast probability of occurrence for event i,
o_i = 1 if the event was observed to occur and 0 if it did not.
Ranges from 0 to 1, with 0 being a perfect forecast.

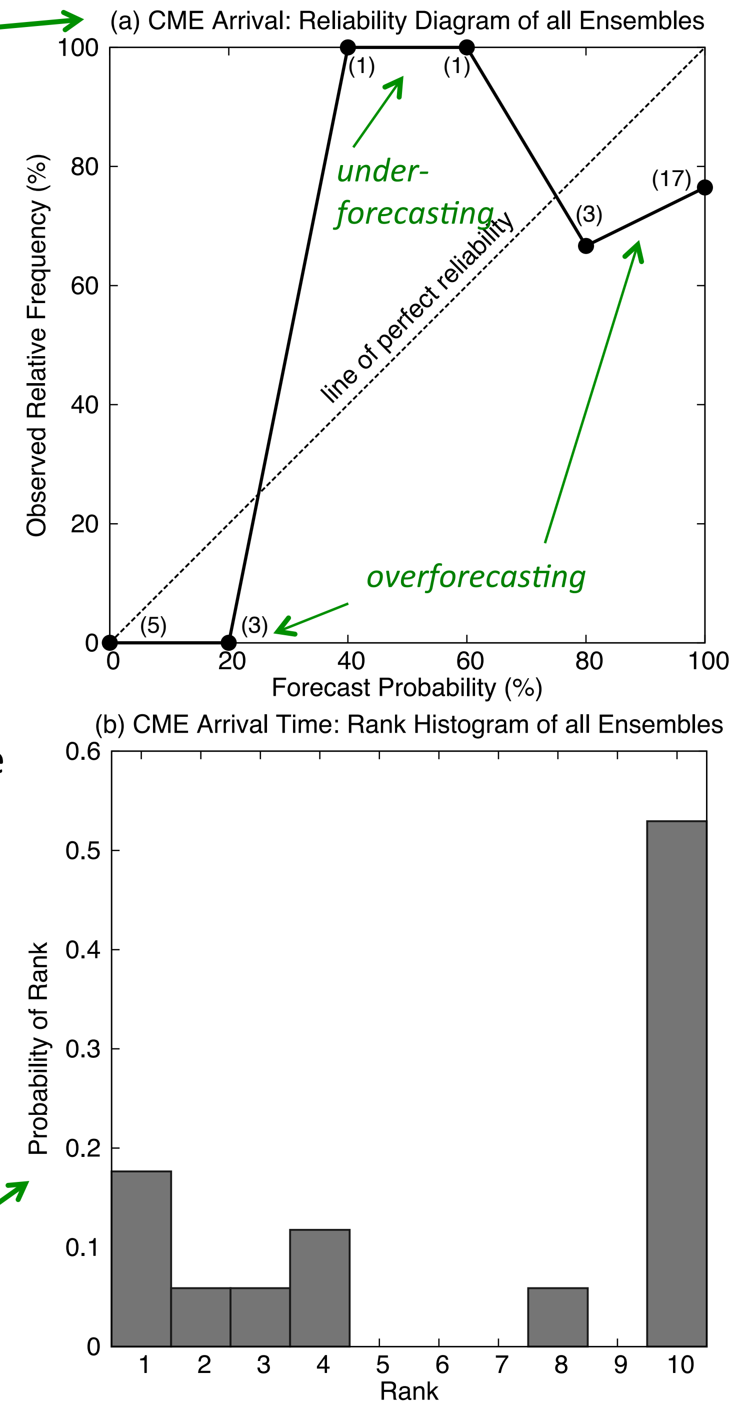
- For CME arrival prediction, the ``event'' here is taken as the ``CME arrival'' and for each event i, p_i is $n_{\text{predicted hits}}/n_{\text{total}}$ (Table 1, column 11).
- The Brier Score computed from all 30 ensemble CME arrival probabilities is 0.15, which indicates that in this sample, on average, the probability p of the CME arriving is fairly accurate.
- However, verification scores reduce the problem to a single measure which can only consider one dimension of the system. Let's examine Reliability.

- Reliable forecasts are those where the observed frequencies of events are in agreement with the forecast probabilities.
- To evaluate the reliability multiple forecasts must be evaluated because a single probabilistic forecast cannot be simply assessed as “right'' or “wrong'' e.g. if a forecast suggests a 30% chance of CME arrival, and the CME does arrive, the forecast is not clearly either “right'' or “wrong''. Therefore, to provide forecast verification for a p=30% chance of CME arrival one would need to compile the statistics of observed CME arrivals for a set of forecasts that predicted a 30% chance of arrival.
- This method is used to construct a reliability diagram can be constructed to determine how well the predicted probabilities of an event correspond to their observed frequencies.

Ib) Likelihood of CME arrival forecast verification: Reliability

The reliability diagram for this sample shows that ensemble modeling is underforecasting in the forecast bins between 20-80%, and slightly overforecasting in the 1-20% and 80-100% forecast bins.

Overforecasting is when the forecast chance of CME arrival is higher than is actually observed; i.e., the CME is observed to arrive less often than is predicted.



How well does the ensemble spread represent the true variability of the observations?

- The observed CME arrival was within the spread of ensemble arrival time predictions for 8 out of 17 of the ensemble runs containing hits.
- For a reliable forecast the set of ensemble member forecast values for a given event and observations should be considered as random samples from the same probability distribution.
- If an n member ensemble and the observation are sorted from earliest to latest arrival times, the observation is equally likely to occur in each of the n+1 possible "ranks". Therefore a histogram of the rank of the observation, "rank histogram", tallied over many events should show be uniform (flat).
- The U-shaped rank histogram for this sample suggests undervariability, indicating that these ensembles to not sample a wide enough spread in CME input parameters.

Summary of Ensemble Kp Predictions

Table 3. Summary of K_P prediction results for 17 ensemble runs containing hits. Columns 1-2: CME start date and time. Columns 3-11: overall probabilistic K_P forecast for each K_P bin assuming equal likelihood of three clock angle scenarios. Underlined K_P probabilities indicate that the NOAA real-time K_P observation falls in this bin and the observed definitive K_P is in column 13. The mean predicted K_P is listed in column 12, along with the overall predicted K_P spread (using plus or minus notation). The Brier Score BS is calculated for each K_P bin and listed on the last line of the table. The Dst sudden storm commencement and minimum values are listed in the last two columns.

CME Onset		Binned Probabilistic K_P Forecast (%)									Mean K_P	Obs.	Dst (nT)	
Date	Time (UT)	1	2	3	4	5	6	7	8	9	& spread	K_P	SSC	min.
2013-01-16	19:00	0	13	$\underline{26}$	28	6	11	9	4	4	4^{+5}_{-2}	4-	+6	-34
2013-04-11	07:24	0	0	$\underline{0}$	0	0	33	5	62	0	7^{+1}_{-1}	3+	+21	-7
2013-06-21	03:12	0	0	0	$\underline{0}$	4	16	23	43	15	7^{+2}_{-2}	5+	....	-49
2013-08-30	02:48	0	0	$\underline{6}$	31	28	33	2	0	0	4^{+3}_{-1}	3+	....	-31
2013-09-29	20:40	0	0	6	26	24	$\underline{39}$	5	0	0	5^{+2}_{-2}	8-	+30	-67
2013-10-06	14:39	0	67	33	0	$\underline{0}$	0	0	0	0	2^{+1}_{-0}	6-	+21	-65
2014-01-07	18:24	0	0	$\underline{0}$	6	8	19	25	26	16	7^{+2}_{-3}	3-	+2	-14
2014-01-30	16:24	0	$\underline{0}$	15	13	33	13	18	8	0	5^{+3}_{-2}	2+	+15	-7
2014-02-12	05:39	0	0	12	25	$\underline{40}$	24	0	0	0	4^{+2}_{-1}	5o	+52	-16
2014-02-18	01:25	0	1	10	21	29	$\underline{26}$	10	2	0	5^{+3}_{-3}	6o	....	-86
2014-02-19	16:00	0	2	30	$\underline{34}$	28	5	0	0	0	4^{+2}_{-2}	4+	+4	-56
2014-02-25	01:09	0	0	1	11	16	$\underline{21}$	22	21	9	6^{+3}_{-3}	5+	....	-99
2014-03-23	03:48	0	0	16	$\underline{28}$	28	24	4	0	0	4^{+3}_{-1}	4-	+20	-18
2014-04-02	13:36	0	0	21	$\underline{19}$	40	12	7	0	0	4^{+3}_{-1}	4o	+16	-16
2014-04-18	13:09	0	0	0	3	$\underline{27}$	17	40	14	0	6^{+2}_{-2}	5o	+25	-24
2014-06-04	15:48	0	0	18	29	36	$\underline{17}$	0	0	0	4^{+2}_{-1}	6+	+31	-38
2014-06-19	17:12	0	0	$\underline{8}$	31	33	25	3	0	0	4^{+3}_{-1}	3o	+14	-9
Brier Score		0.00	0.09	0.27	0.19	0.17	0.17	0.02	0.04	0.00				

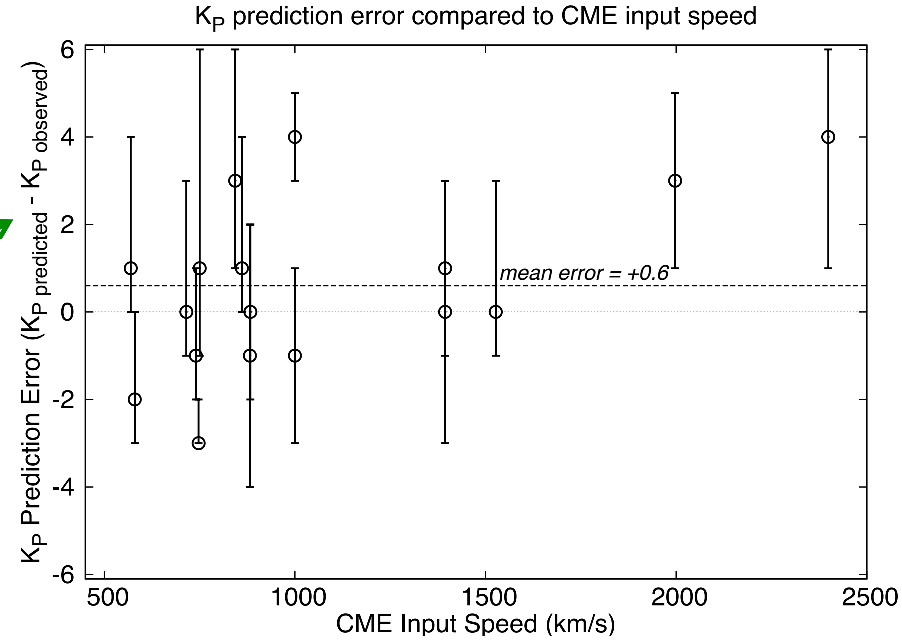
II) Kp forecast verification

- We compute the prediction error

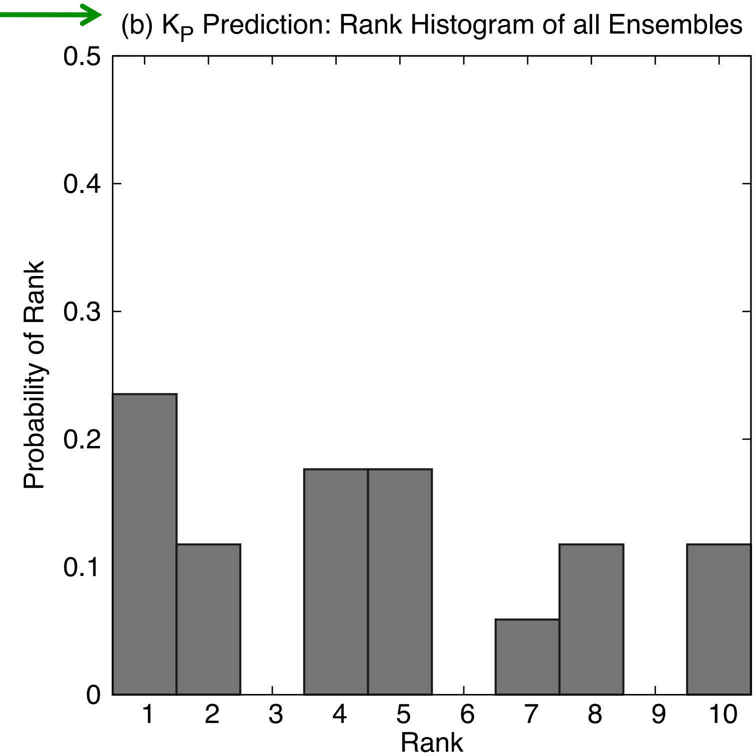
$$Kp_{error} = Kp_{predicted} - Kp_{observed}$$

using the mean ensemble predicted Kp and find: MAE=1.5, RMSE=2.0, and ME=+0.6

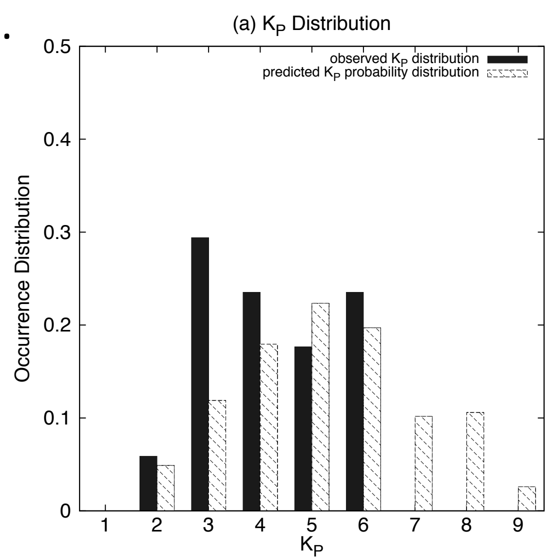
- Plot of Kp prediction error vs. speed shows that Kp is usually overpredicted when CME input speeds are above 800-1000 km/s. *This is likely due to an overestimation of the CME dynamic pressure at Earth by the model.*



- The rank histogram for the Kp prediction has an overall flat shape, with more occurrence at rank 1 (the observed Kp was less than the predicted range) and less occurrence in the higher ranks which shows the bias for Kp overprediction.



Note, that the rank histogram does not indicate how ``good'' forecasts are but only measures whether the observed probability distribution is well represented by the ensemble.



Summary

• The ensemble prediction approach provides a probabilistic forecast which includes an estimation of arrival time uncertainty from the spread in predictions and a forecast confidence in the likelihood of CME arrival.

• First results give a mean absolute arrival time error of 12.3 hours, RMSE of 13.9 hours, and mean error of -5.8 hours (early bias), based on a sample of 30 CME events for which ensemble simulations were performed. The ensemble mean absolute error and RMSE are both comparable with other CME arrival time prediction errors reported in the literature.

• It was found that the correct rejection rate is 62%, the false-alarm rate is 38%, correct alarm ratio is 77%, and false alarm ratio is 23%.

•The Brier Score of 0.15 for all 30 ensemble CME arrival probabilities indicates that in this sample, on average, the predicted probability of the CME arriving is fairly accurate.

• However, the reliability diagram shows that the ensemble simulations are underforecasting the likelihood that the CME will arrive in the forecast bins between 20-80%, and slightly overforecasting in the 1-20% and 80-100% forecast bins.

• For 8 out of 17 of the ensemble runs containing hits, the observed CME arrival was within the spread of ensemble arrival time predictions. The initial distribution of CME input parameters was shown to be an important influence on the accuracy of CME arrival time predictions. The rank histogram suggests undervariability in initial conditions; i.e., these ensembles do not sample a wide enough spread in CME input parameters.

• The observed Kp was within ±1 of the predicted mean Kp for 11 out of 17 of the ensembles.

• The Kp prediction errors computed from the mean predicted Kp show a mean absolute error of 1.4, RMSE of 1.8, and mean error +0.4.

• There is a known overall tendency for the overprediction of Kp, generally found for CME input speeds above 800-1000 km/s.